

## **O AEROPORTO FRANCISCO SÁ CARNEIRO E A SUA RELAÇÃO COM O TURISMO NA REGIÃO NORTE DE PORTUGAL**

## **THE FRANCISCO SÁ CARNEIRO AIRPORT AND ITS RELATION WITH TOURISM IN THE NORTH REGION OF PORTUGAL**

**Helena de Barros<sup>1</sup>, Hugo Alonso<sup>2</sup>**

<sup>1</sup>Universidade Lusófona do Porto, Portugal, [hcoelhobarros@gmail.com](mailto:hcoelhobarros@gmail.com)

<sup>2</sup>Universidade Lusófona do Porto, Universidade de Aveiro, CIDMA, Portugal,  
[hugo.alonso@ulp.pt](mailto:hugo.alonso@ulp.pt)

# **O AEROPORTO FRANCISCO SÁ CARNEIRO E A SUA RELAÇÃO COM O TURISMO NA REGIÃO NORTE DE PORTUGAL**

## **THE FRANCISCO SÁ CARNEIRO AIRPORT AND ITS RELATION WITH TOURISM IN THE NORTH REGION OF PORTUGAL**

### **RESUMO**

O Aeroporto Francisco Sá Carneiro e o turismo na região Norte de Portugal têm crescido de forma muito significativa nos últimos anos. Este artigo apresenta um estudo da relação entre o número de passageiros que circulam no aeroporto e as dormidas na região. O estudo é baseado numa análise de regressão linear simples. Como resultado, propõe-se uma forma de prever as dormidas, conhecido o número de passageiros.

***Palavras-chave:** Aeroporto Francisco Sá Carneiro; análise de regressão; região Norte de Portugal; turismo.*

### **ABSTRACT**

The Francisco Sá Carneiro Airport and tourism in the North region of Portugal have grown very significantly in recent years. This paper presents a study of the relation between the number of passengers traveling at the airport and the number of overnight stays in the region. The study is based on a simple linear regression analysis. As a result, it is proposed a way to predict overnight stays, known the number of passengers.

***Keywords:** Francisco Sá Carneiro Airport; North region of Portugal; regression analysis; tourism.*

## 1. INTRODUÇÃO

O Aeroporto Francisco Sá Carneiro, também conhecido por Aeroporto do Porto, é o maior aeroporto da região Norte de Portugal e o segundo maior do país. Como é do conhecimento geral, este aeroporto tem recebido várias distinções. A liberalização do transporte aéreo no final do século passado e o grande investimento na expansão e modernização das suas instalações em anos recentes contribuíram de forma significativa para o seu crescimento.

A par do crescimento do Aeroporto do Porto, tem-se assistido a um desenvolvimento progressivo do turismo na região Norte de Portugal. Esta tendência surge de forma natural, já que o transporte aéreo e o turismo têm uma forte relação entre si. O desenvolvimento do turismo na região é evidente, por exemplo, no aumento do número de dormidas, e tem sido reconhecido além-fronteiras, em particular, com as várias distinções que o Porto tem recebido como melhor destino europeu.

O principal objetivo deste trabalho é apresentar um estudo da relação entre o número de passageiros que circulam no Aeroporto do Porto e as dormidas no Norte de Portugal. O estudo é baseado numa análise de regressão linear simples. Tanto quanto se sabe, não há nada similar na literatura. O mais próximo que se encontrou foi (Costa, 2016). Neste artigo, a autora apresenta um estudo de correlação, mas não de regressão, ou seja, mede a associação entre as duas variáveis, mas não estabelece uma fórmula que as relacione e que permita prever uma delas à custa da outra. Além disso, a autora considera dados anuais de 2003 a 2013, enquanto que, neste trabalho, são considerados dados anuais de 1999 a 2017. Como resultado da análise de regressão realizada, foi possível concluirmos que existe uma relação linear estatisticamente significativa entre a variação anual do número de passageiros que circulam no Aeroporto do Porto e a correspondente variação anual das dormidas no Norte de Portugal. A fórmula que obtivemos permite prever as dormidas, conhecido o número de passageiros.

O resto deste artigo está dividido em duas partes. A primeira parte mostra o estudo da relação entre o número de passageiros que circulam no Aeroporto do Porto e as dormidas no Norte de Portugal e a segunda parte dá conta das conclusões e do trabalho futuro.

## 2. ESTUDO DA RELAÇÃO ENTRE O NÚMERO DE PASSAGEIROS QUE CIRCULAM NO AEROPORTO FRANCISCO SÁ CARNEIRO E AS DORMIDAS NA REGIÃO NORTE DE PORTUGAL

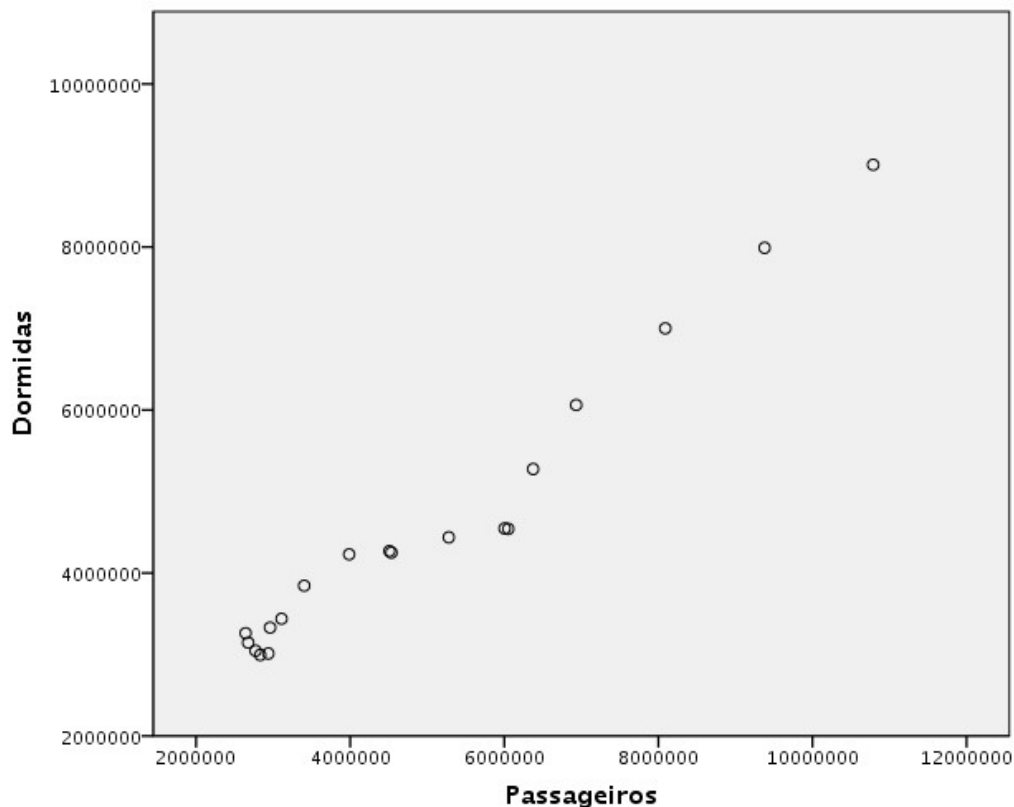
O estudo da relação entre o número de passageiros que circulam no Aeroporto Francisco Sá Carneiro (AFSC) e as dormidas na região Norte de Portugal baseia-se na seguinte amostra de 19 observações anuais, referentes ao período de 1999 a 2017, destas duas variáveis:

Ano	Passageiros no AFSC	Dormidas na região Norte
1999	2 832 722	2 994 353
2000	2 937 582	3 012 673
2001	2 771 169	3 046 000
2002	2 642 420	3 262 430
2003	2 675 823	3 145 780
2004	2 960 553	3 330 650
2005	3 108 271	3 438 700
2006	3 402 763	3 844 500
2007	3 986 860	4 228 965
2008	4 534 829	4 250 764
2009	4 508 533	4 269 967
2010	5 279 716	4 437 756
2011	6 004 500	4 547 005
2012	6 051 081	4 541 919
2013	6 374 045	5 276 137
2014	6 932 614	6 061 742
2015	8 088 907	7 001 899
2016	9 378 206	7 989 922
2017	10 787 630	9 008 846

*Quadro 1 - Número anual de passageiros que circulam no Aeroporto Francisco Sá Carneiro e dormidas anuais na região Norte de Portugal*

Fonte: Adaptado pelo autor de (INE, 2000a a 2017a, 2018, 2000b a 2017b; ANA, 2018).

O diagrama de dispersão dos dados é o seguinte:



*Figura 1 - Dormidas na região Norte vs. número de passageiros que circulam no AFSC.*

Fonte: O autor.

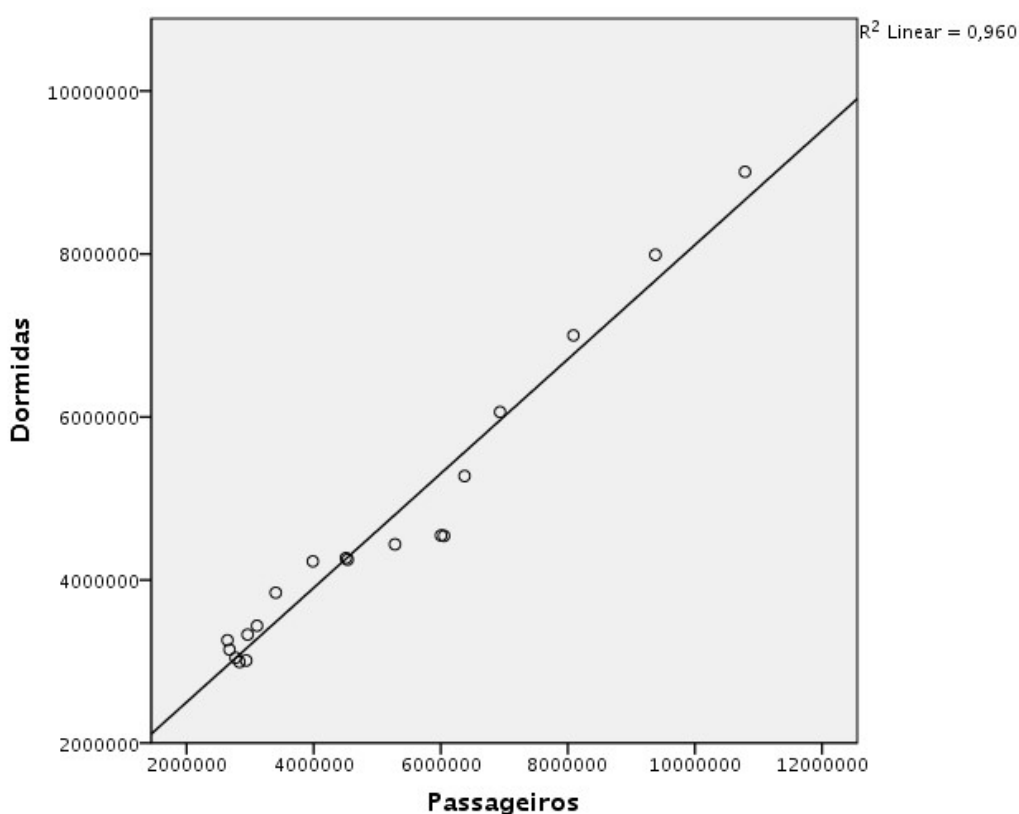
Observando a figura, é possível concluir que, na amostra, quanto maior é o número de passageiros que circulam no aeroporto, maior tende a ser o número de dormidas na região Norte. Além disso, na amostra, a relação entre as duas variáveis parece ser aproximadamente linear. Esta conclusão preliminar motivou a escolha do modelo de regressão linear simples para representar a relação entre as duas variáveis:

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (1)$$

onde  $y$  é o número de dormidas na região Norte e  $x$  é o número de passageiros que circulam no aeroporto. Usando os dados da amostra, obteve-se através do SPSS a seguinte equação de regressão estimada pelo método dos mínimos quadrados:

$$\hat{y} = 1099048,42 + 0,701x \quad (2)$$

A próxima figura mostra esta equação no diagrama de dispersão. O coeficiente de determinação é  $r^2 = 0,960$ , ou seja, 96,0% da variabilidade registada na amostra para o número de dormidas na região Norte é explicada pela equação de regressão estimada. O coeficiente de correlação linear de Pearson é  $r = 0,980$ . O facto de  $r^2$  e  $r$  serem elevados (próximos de 1) significa que a equação de regressão estimada (2) se ajusta muito bem aos dados da amostra. No entanto, a equação (2) não deve ser usada para prever o número de dormidas na região Norte, conhecido o número de passageiros que circulam no aeroporto, se não for demonstrada a significância estatística da relação (1) que se assumiu entre as duas variáveis. Os testes de significância são baseados em hipóteses sobre o erro  $\epsilon$  do modelo de regressão (1), ou seja, só devem ser aplicados se esses pressupostos se verificarem (Anderson et al., 2011). Como tal, segue-se um estudo da validade das hipóteses sobre o erro  $\epsilon$ .



*Figura 2 - Reta de regressão estimada das dormidas na região Norte em função do número de passageiros que circulam no AFSC.*

Fonte: O autor.

O procedimento usado para investigar a validade das hipóteses sobre o erro  $\varepsilon$  do modelo de regressão (1) é conhecido por análise residual (Anderson et al., 2011). Note-se que, a partir de (1), o erro é dado por

$$\varepsilon = y - (\beta_0 + \beta_1 x) \quad (3)$$

mas que não pode ser calculado nos casos da amostra que se está a considerar, porque  $\beta_0$  e  $\beta_1$  são desconhecidos. No entanto, a partir da equação de regressão estimada (2), tem-se que  $\beta_0 \approx b_0 = 1099048,42$  e  $\beta_1 \approx b_1 = 0,701$  e, portanto,

$$\varepsilon \approx e = y - (1099048,42 + 0,701x) \quad (4)$$

onde  $e$  é designado por resíduo. Note-se que o resíduo  $e$  pode ser calculado nos casos da amostra que se está a considerar e que, embora aproximada, é única a informação disponível sobre o erro  $\varepsilon$ . Assim, o estudo da validade das hipóteses sobre o erro  $\varepsilon$  é baseado na análise dos resíduos calculados.

### **Hipótese 1: O erro $\varepsilon$ é uma variável aleatória normalmente distribuída**

Fixemos o nível de significância em  $\alpha=0,01$  e consideremos as hipóteses seguintes:

$$\begin{cases} H_0: \text{os resíduos têm distribuição normal} \\ H_1: \text{os resíduos não têm distribuição normal} \end{cases}$$

De acordo com (Hall et al., 2011), o teste de normalidade de Shapiro-Wilk é o mais indicado para amostras de pequenas dimensões, ou seja, onde o número de observações é  $n \leq 30$ , como é o caso aqui, em que  $n = 19$ . Usando o SPSS, verifica-se que o nível de significância observado neste teste é Sig.=0,011. Logo, como se trata de um valor superior ao fixado ( $\alpha=0,01$ ), não se rejeita a hipótese  $H_0$  de os resíduos terem uma distribuição normal.

### **Hipótese 2: O erro $\epsilon$ é uma variável aleatória de média 0**

Fixemos o nível de significância em  $\alpha=0,01$  e consideremos as hipóteses seguintes:

$$\begin{cases} H_0: \text{os resíduos têm média } 0 \\ H_1: \text{os resíduos não têm média } 0 \end{cases}$$

A aplicação no SPSS de um teste  $t$  para a média de uma população normal com variância desconhecida (Hall et al., 2011) leva a um nível de significância observado de Sig.=1. Logo, como se trata de um valor superior ao fixado ( $\alpha=0,01$ ), não se rejeita a hipótese  $H_0$  de os resíduos terem média 0.

### **Hipótese 3: A variância de $\epsilon$ , representada por $\sigma^2$ , é a mesma para todos os valores de $x$**

Fixemos o nível de significância em  $\alpha=0,01$  e consideremos as hipóteses seguintes:

$$\begin{cases} H_0: \text{Os resíduos têm a mesma variância (homoscedasticidade)} \\ H_1: \text{Os resíduos não têm a mesma variância (heteroscedasticidade)} \end{cases}$$

A homoscedasticidade pode ser testada de forma adequada com recurso ao teste de White (Murteira et al., 2015). A aplicação no SPSS deste teste conduz a um nível de significância observado de Sig.=0,15. Logo, como se trata de um valor superior ao fixado ( $\alpha=0,01$ ), não se rejeita a hipótese  $H_0$  de os resíduos terem a mesma variância (homoscedasticidade).

### **Hipótese 4: Os valores de $\epsilon$ são independentes**

Fixemos o nível de significância em  $\alpha=0,01$  e consideremos as hipóteses seguintes:

$$\begin{cases} H_0: \text{Os resíduos não estão autocorrelacionados (independência)} \\ H_1: \text{Os resíduos estão autocorrelacionados (dependência)} \end{cases}$$

A deteção de autocorrelação (dependência) pode ser feita com o teste de Durbin-Watson (Murteira et al., 2015). A existir autocorrelação, ela pode ser positiva ou negativa. Se os



resíduos estiverem positivamente autocorrelacionados, verifica-se que cada resíduo tende a manter o sinal do resíduo anterior. Deste modo, há vários resíduos seguidos com sinal positivo e vários resíduos seguidos com sinal negativo. Se a autocorrelação dos resíduos for negativa, então cada resíduo tende a ter o sinal contrário do resíduo anterior. A estatística de teste é representada por  $d$  e a ideia é compará-la com valores críticos  $d_L$  e  $d_U$ :

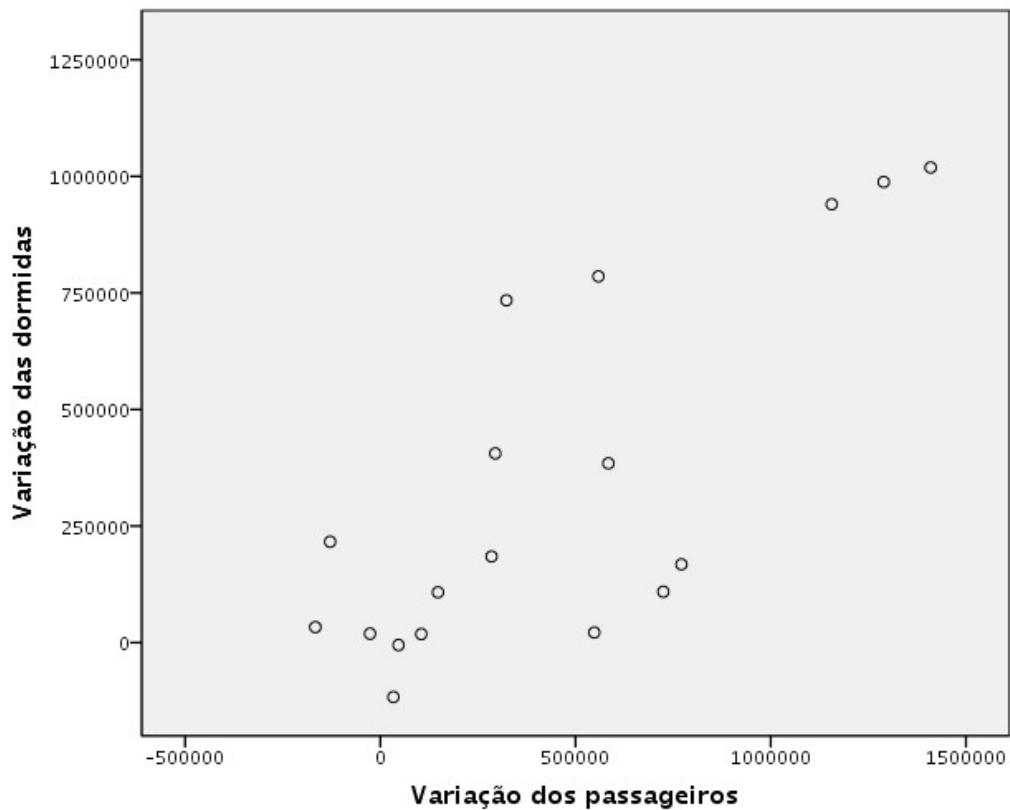
- Se  $d < d_L$  há evidência de autocorrelação positiva e se  $4 - d < d_L$  há evidência de autocorrelação negativa;
- Se  $d > d_U$  não há evidência de autocorrelação positiva e se  $4 - d > d_U$  não há evidência de autocorrelação negativa;
- Se  $d_L < d < d_U$  e  $d_L < 4 - d < d_U$ , então não é possível concluir nada.

Usando o SPSS, obtém-se  $d = 0,484$ . Logo,  $4 - d = 3,516$ . Além disso, consultando (Savin & White, 1977), tem-se que  $d_L = 0,928$  e  $d_U = 1,133$ , tendo em conta que o nível de significância fixado é  $\alpha=0,01$ , o número de variáveis independentes é  $k = 1$  e o número de observações é  $n = 19$ . Assim, como  $d < d_L$ , há evidência de autocorrelação positiva e, como  $4 - d > d_U$ , não há evidência de autocorrelação negativa. Em suma, o teste rejeita a hipótese  $H_0$  e aceita  $H_1$ , ou seja, conclui-se que os resíduos não são independentes.

**Conclusão:** Como uma das hipóteses sobre o erro  $\varepsilon$  não se verifica, seriam questionáveis os resultados da aplicação dos testes de significância ao estudo da significância estatística da relação (1) que se assumiu entre o número de passageiros que circulam no Aeroporto Francisco Sá Carneiro e as dormidas na região Norte de Portugal. Assim, não é possível garantir que tal relação seja apropriada para prever as dormidas, conhecido o número de passageiros.

Em face da conclusão anterior, decidiu-se seguir uma nova abordagem: estudar a relação entre a variação anual do número de passageiros que circulam no Aeroporto Francisco Sá Carneiro e a correspondente variação anual das dormidas na região Norte de Portugal. Este novo estudo baseia-se numa amostra de 18 observações destas duas novas variáveis, obtida a partir da amostra do estudo anterior (ver Quadro 1). Por exemplo, de 1999 para 2000, o número de passageiros passou de 2 832 722 para 2 937 582, ou seja, registou-se uma variação de 104 860 passageiros e o número de dormidas passou de 2 994 353 para 3 012 673, isto é,

verificou-se uma variação de 18 320 dormidas. O diagrama de dispersão dos novos dados (variações) é o seguinte:



*Figura 3 - Variação anual das dormidas na região Norte vs. variação anual do número de passageiros que circulam no AFSC.*

Fonte: O autor.

Comparando esta figura com a correspondente do estudo anterior (Figura 1), é possível concluir que, na nova amostra, os dados não estão tão dispostos segundo uma linha reta oblíqua. Ainda assim, escolheu-se o modelo de regressão linear simples para representar a relação entre as duas novas variáveis:

$$\Delta y = \beta_0 + \beta_1 \Delta x + \varepsilon \quad (5)$$

onde  $\Delta y$  é a variação anual do número de dormidas na região Norte e  $\Delta x$  é a variação anual do número de passageiros que circulam no aeroporto. Usando os dados da nova amostra, obteve-se através do SPSS a seguinte equação de regressão estimada pelo método dos mínimos quadrados:

$$\Delta\hat{y} = 60592,483 + 0,6194x. \quad (6)$$

A próxima figura mostra esta equação no diagrama de dispersão. O coeficiente de determinação é  $r^2 = 0,596$ , ou seja, 59,6% da variabilidade registrada na amostra para a variação anual do número de dormidas na região Norte é explicada pela equação de regressão estimada. O coeficiente de correlação linear de Pearson é  $r = 0,772$ . Apesar de  $r^2$  e  $r$  não serem tão elevados como no estudo anterior, são, ainda assim, bastante razoáveis. No entanto, a equação (6) não deve ser usada para prever a variação anual do número de dormidas na região Norte, conhecida a variação anual do número de passageiros que circulam no aeroporto, se não for demonstrada a significância estatística da relação (5) que se assumiu entre as duas novas variáveis. Conforme já foi referido, os testes de significância são baseados em hipóteses sobre o erro  $\varepsilon$ , neste caso do modelo de regressão (5), ou seja, só devem ser aplicados se esses pressupostos se verificarem. Como tal, segue-se um estudo da validade das hipóteses sobre o erro  $\varepsilon$ , baseado na análise dos resíduos que podem ser calculados nos casos da nova amostra de acordo com a fórmula

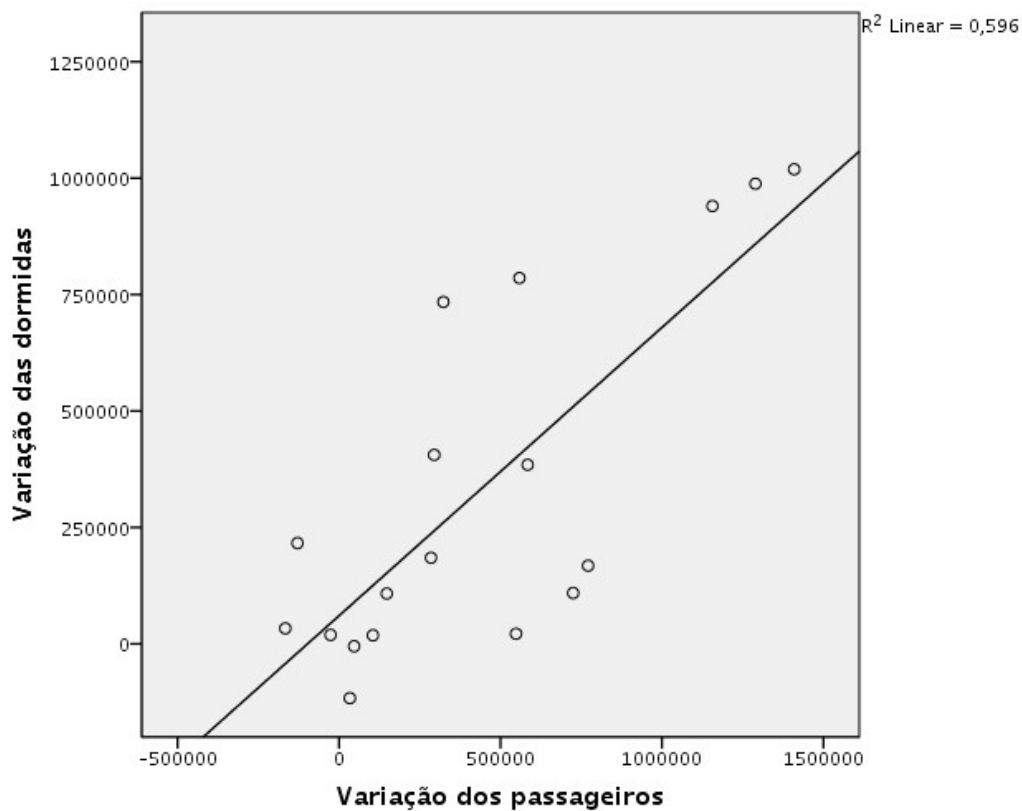
$$e = \Delta y - (60592,483 + 0,6194x). \quad (7)$$

### **Hipótese 1: O erro $\varepsilon$ é uma variável aleatória normalmente distribuída**

Fixemos o nível de significância em  $\alpha=0,01$  e consideremos as hipóteses seguintes:

$$\begin{cases} H_0: \text{os resíduos têm distribuição normal} \\ H_1: \text{os resíduos não têm distribuição normal} \end{cases}$$

Usando o SPSS para aplicar o teste de normalidade de Shapiro-Wilk, obtém-se um nível de significância observado de Sig.=0,688. Logo, como se trata de um valor superior ao fixado ( $\alpha=0,01$ ), não se rejeita a hipótese  $H_0$  de os resíduos terem uma distribuição normal.



*Figura 4 - Reta de regressão estimada da variação anual das dormidas na região Norte em função da variação anual do número de passageiros que circulam no AFSC.*

Fonte: O autor.

## **Hipótese 2: O erro $\epsilon$ é uma variável aleatória de média 0**

Fixemos o nível de significância em  $\alpha=0,01$  e consideremos as hipóteses seguintes:

$$\begin{cases} H_0: \text{os resíduos têm média } 0 \\ H_1: \text{os resíduos não têm média } 0 \end{cases}$$

A aplicação no SPSS de um teste  $t$  para a média de uma população normal com variância desconhecida leva a um nível de significância observado de Sig.=1. Logo, como se trata de um valor superior ao fixado ( $\alpha=0,01$ ), não se rejeita a hipótese  $H_0$  de os resíduos terem média 0.

**Hipótese 3: A variância de  $\varepsilon$ , representada por  $\sigma^2$ , é a mesma para todos os valores de  $\Delta x$**

Fixemos o nível de significância em  $\alpha=0,01$  e consideremos as hipóteses seguintes:

$$\begin{cases} H_0: \text{Os resíduos têm a mesma variância (homoscedasticidade)} \\ H_1: \text{Os resíduos não têm a mesma variância (heteroscedasticidade)} \end{cases}$$

A aplicação no SPSS do teste de White conduz a um nível de significância observado de Sig.=0,056. Logo, como se trata de um valor superior ao fixado ( $\alpha=0,01$ ), não se rejeita a hipótese  $H_0$  de os resíduos terem a mesma variância (homoscedasticidade).

**Hipótese 4: Os valores de  $\varepsilon$  são independentes**

Fixemos o nível de significância em  $\alpha=0,01$  e consideremos as hipóteses seguintes:

$$\begin{cases} H_0: \text{Os resíduos não estão autocorrelacionados (independência)} \\ H_1: \text{Os resíduos estão autocorrelacionados (dependência)} \end{cases}$$

A deteção de autocorrelação (dependência) é feita com o teste de Durbin-Watson. A estatística de teste é representada por  $d$  e, tal como no estudo anterior, a ideia é compará-la com valores críticos  $d_L$  e  $d_U$ . Usando o SPSS, obtém-se  $d = 1,171$ . Logo,  $4 - d = 2,829$ . Além disso, consultando (Savin & White, 1977), tem-se que  $d_L = 0,902$  e  $d_U = 1,118$ , uma vez que o nível de significância fixado é  $\alpha=0,01$ , o número de variáveis independentes é  $k = 1$  e o número de observações é  $n = 18$ . Assim, como  $d > d_U$ , não há evidência de autocorrelação positiva e, como  $4 - d > d_U$ , também não há evidência de autocorrelação negativa. Deste modo, o teste não rejeita a hipótese  $H_0$  de os resíduos serem independentes.

**Conclusão:** Como todas as hipóteses sobre o erro  $\varepsilon$  se verificam, faz sentido aplicar testes de significância ao estudo da significância estatística da relação (5) que se assumiu entre a variação anual do número de passageiros que circulam no Aeroporto Francisco Sá Carneiro e a correspondente variação anual das dormidas na região Norte de Portugal.

## Teste de significância

Consideremos, então, a relação (5):  $\Delta y = \beta_0 + \beta_1 \Delta x + \varepsilon$ , onde  $\Delta y$  é a variação anual do número de dormidas na região Norte e  $\Delta x$  é a variação anual do número de passageiros que circulam no aeroporto. Fixemos o nível de significância em  $\alpha=0,01$  e consideremos as hipóteses seguintes:

$$\begin{cases} H_0: \beta_1 = 0 \\ H_1: \beta_1 \neq 0 \end{cases}$$

Usando o SPSS para aplicar o teste  $t$ , obtém-se um nível de significância observado de  $\text{Sig.}=0,000$ . Como este é inferior ao fixado ( $\alpha=0,01$ ), rejeita-se a hipótese  $H_0$  e aceita-se  $H_1$ , ou seja, conclui-se que  $\beta_1 \neq 0$ . Um intervalo de confiança a 99% para  $\beta_1$  é  $]0,247; 0,991[$  e, tal como esperado, 0 não pertence ao intervalo. Assim, a relação linear entre as duas variáveis  $\Delta y$  e  $\Delta x$  é estatisticamente significativa.

**Consequência:** demonstrada a significância estatística do modelo de regressão (5),

$$\Delta y = \beta_0 + \beta_1 \Delta x + \varepsilon, \quad (8)$$

pode-se usar a equação de regressão estimada (6),

$$\Delta \hat{y} = 60592,483 + 0,619 \Delta x, \quad (9)$$

para prever a variação anual do número de dormidas na região Norte, conhecida a variação anual do número de passageiros que circulam no aeroporto. De acordo com esta equação, estima-se um aumento de 619 unidades em  $\Delta y$  por cada aumento de 1 000 unidades em  $\Delta x$ , ou seja, um aumento de 619 dormidas na variação anual das dormidas por cada aumento de 1000 passageiros na variação anual do número de passageiros.

Mesmo no final da escrita deste artigo, teve-se acesso a um valor aproximado do número de passageiros que circularam no aeroporto em 2018, a saber: 11 939 000 (VINCI, 2019). Tendo em conta que o número observado no ano anterior foi igual a 10 787 630,

registou-se uma variação de  $\Delta x = 1\,151\,370$  passageiros de 2017 para 2018. Este valor está compreendido entre o menor valor e o maior valor registados na amostra a partir da qual foi obtida a equação (6). Usando esta equação, estima-se que a variação do número de dormidas na região Norte de 2017 para 2018 tenha sido

$$\Delta \hat{y} = 60\,592,483 + 0,619 \times 1\,151\,370 \approx 773\,291. \quad (10)$$

Usando o SPSS, obteve-se

$$\Delta y \in ] - 26\,714; 1\,573\,220[ \quad (11)$$

para intervalo de previsão individual, com um nível de confiança de 99%. Como o número de dormidas em 2017 foi de 9 008 846, é possível dizer que uma estimativa do número de dormidas em 2018 é de

$$\hat{y} = 9\,008\,846 + 773\,291 = 9\,782\,137 \quad (12)$$

e que, com 99% de confiança,

$$y \in ]9\,008\,846 - 26\,714; 9\,008\,846 + 1\,573\,220[ = ]8\,982\,132; 10\,582\,066[. \quad (13)$$

Infelizmente, na altura da escrita deste texto, não era ainda conhecido o valor das dormidas em 2018 e, portanto, não foi possível compará-lo com a estimativa fornecida e perceber se estava ou não no intervalo indicado.

### 3. CONCLUSÕES E TRABALHO FUTURO

O transporte aéreo e o turismo têm uma forte relação entre si e têm vindo a crescer de ano para ano. Um fator importante para o elevado crescimento do número de turistas no mundo foi sem dúvida a chegada das companhias aéreas de baixo custo.

O Aeroporto Francisco Sá Carneiro e o turismo na região Norte de Portugal têm mostrado um desenvolvimento progressivo e conquistado vários prémios e distinções nos últimos anos. Este artigo considerou o problema de estudar a relação entre o número de

passageiros que circulam no aeroporto e as dormidas na região. Foram recolhidos dados anuais destas duas variáveis, tendo-se obtido uma amostra referente ao período de 1999 a 2017. Como resultado da análise de regressão realizada, foi possível concluir que existe uma relação linear estatisticamente significativa entre a variação anual do número de passageiros que circulam no Aeroporto Francisco Sá Carneiro e a correspondente variação anual das dormidas na região Norte de Portugal. Neste contexto, estima-se que haja um aumento de 619 dormidas na variação anual das dormidas por cada aumento de 1 000 passageiros na variação anual do número de passageiros. As fórmulas obtidas permitem prever as dormidas, conhecido o número de passageiros, tal como foi exemplificado para o ano de 2018.

Relativamente às perspetivas de trabalho futuro, seria interessante aprofundar o estudo realizado aqui, considerando mais dados, por exemplo, mensais, e eventualmente seguindo uma análise de regressão múltipla, onde a previsão das dormidas num dado instante se fizesse à custa das dormidas e do número de passageiros em instantes anteriores. Seria igualmente interessante e relevante estudar a relação entre o número de passageiros que circulam no Aeroporto Francisco Sá Carneiro e os proveitos globais do turismo para a região do Norte de Portugal, procurando perceber, assim, o impacto financeiro que o aeroporto tem no turismo da região.

## **AGRADECIMENTOS**

Este trabalho foi parcialmente financiado pela Fundação para a Ciência e Tecnologia (FCT), através do Centro de Investigação e Desenvolvimento em Matemática e Aplicações (CIDMA), no âmbito do projeto UID/MAT/04106/2019.

## **REFERÊNCIAS**

Aeroportos e Navegação Aérea [ANA] (2018). “Relatório de Gestão e Contas 2017”. Obtido de [https://www.ana.pt/sites/default/files/documents/relatorio\\_gestao\\_contas\\_2017\\_pt\\_0.pdf](https://www.ana.pt/sites/default/files/documents/relatorio_gestao_contas_2017_pt_0.pdf)

Anderson, D. R., Sweeney, D. J., & Williams, T. A. (2011). Statistics for Business and Economics (11th edition). Mason, OH: South-Western.

Costa, V. (2016). “Baking up the development of a peripheric region throught international tourism. The case of Oporto airport and Portugal's northern region”. In L. Halász et al. (Eds.),



Responsible Business & Tourism and the Role of Education at Responsible Thinking (pp. 36-52). Budapest, Hungary: Kodolányi János University of Applied Sciences.

Hall, A., Neves, C., & Pereira, P. (2011). Grande Maratona de Estatística no SPSS. Lisboa, Portugal: Escolar Editora.

Instituto Nacional de Estatística [INE] (2000a). Estatísticas do Turismo 1999. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2000b). Estatísticas dos Transportes e Comunicações 1999. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2001a). Estatísticas do Turismo 2000. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2001b). Estatísticas dos Transportes e Comunicações 2000. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2002a). Estatísticas do Turismo 2001. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2002b). Estatísticas dos Transportes e Comunicações 2001. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2003a). Estatísticas do Turismo 2002. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2003b). Estatísticas dos Transportes e Comunicações 2002. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2004a). Estatísticas do Turismo 2003. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2004b). Estatísticas dos Transportes e Comunicações 2003. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2005a). Estatísticas do Turismo 2004. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2005b). Estatísticas dos Transportes e Comunicações 2004. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2006a). Estatísticas do Turismo 2005. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2006b). Estatísticas dos Transportes e Comunicações 2005. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2007a). Estatísticas do Turismo 2006. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2007b). Estatísticas dos Transportes e Comunicações 2006. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2008a). Estatísticas do Turismo 2007. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2008b). Estatísticas dos Transportes e Comunicações 2007. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2009a). Estatísticas do Turismo 2008. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2009b). Estatísticas dos Transportes e Comunicações 2008. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2010a). Estatísticas do Turismo 2009. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2010b). Estatísticas dos Transportes e Comunicações 2009. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2011a). Estatísticas do Turismo 2010. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2011b). Estatísticas dos Transportes e Comunicações 2010. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2012a). Estatísticas do Turismo 2011. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2012b). Estatísticas dos Transportes e Comunicações 2011. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2013a). Estatísticas do Turismo 2012. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2013b). Estatísticas dos Transportes e Comunicações 2012. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2014a). Estatísticas do Turismo 2013. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2014b). Estatísticas dos Transportes e Comunicações 2013. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2015a). Estatísticas do Turismo 2014. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2015b). Estatísticas dos Transportes e Comunicações 2014. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2016a). Estatísticas do Turismo 2015. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2016b). Estatísticas dos Transportes e Comunicações 2015. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2017a). Estatísticas do Turismo 2016. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2017b). Estatísticas dos Transportes e Comunicações 2016. Lisboa, Portugal: Instituto Nacional de Estatística.

Instituto Nacional de Estatística [INE] (2018). Estatísticas do Turismo 2017. Lisboa, Portugal: Instituto Nacional de Estatística.

Murteira, B., Ribeiro, C. S., Silva, J. A., & Pimenta, C. (2015). Introdução à Estatística (3ª edição). Lisboa, Portugal: Escolar Editora.

Savin, N. E., & White, K. J. (1977). “The Durbin-Watson Test for Serial Correlation with Extreme Sample Sizes or Many Regressors”. *Econometrica*, 45, 1989-1996.

VINCI (2019). “VINCI Airports – tráfego do quarto trimestre e do ano 2018”. Obtido de [https://www.aeroporto lisboa.pt/pt/system/files/documents/prvinci\\_resultados\\_trafego\\_aeropertos\\_q42018.pdf](https://www.aeroporto lisboa.pt/pt/system/files/documents/prvinci_resultados_trafego_aeropertos_q42018.pdf)